Sequence-specific DNA recognition by cyclic pyrrole-imidazole cysteine-derived polyamide dimers.

AUTHOR(S):

Yamamoto, Makoto; Bando, Toshikazu; Morinaga, Hironobu; Kawamoto, Yusuke; Hashiya, Kaori; Sugiyama, Hiroshi

CITATION:

Yamamoto, Makoto ...[et al]. Sequence-specific DNA recognition by cyclic pyrrole-imidazole cysteine-derived polyamide dimers.. Chemistry - A European Journal 2013, 20(3): 752-759

ISSUE DATE:

2013-12-11

URL:

http://hdl.handle.net/2433/198837

Sequence-Specific DNA Recognition by Dimerized Cysteine Cyclic Pyrrole-Imidazole Polyamides

Makoto Yamamoto[a], Toshikazu Bando* [a], Hironobu Morinaga[a], Yusuke Kawamoto[a], Kaori Hashiya[a] and Hiroshi Sugiyama* [a], [b], [c]

[a] Department of Chemistry, Graduate School of Science Kyoto University, Sakyo, Kyoto 606-8502, Japan

[b] Institute for Integrated Cell-Materials Science (iCeMS) Kyoto University, Sakyo, Kyoto 606-8502, Japan

[c] CREST, Japan Science and Technology Corporation (JST), Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan

Supporting information for this article is available on the WWW under http://xxx.

Abstract: Pyrrole-imidazole (PI) polyamides bind to the minor groove of the DNA duplex in a sequence-specific manner and thus have the potential to regulate gene expression. To date, various types of PI polyamides have been designed as sequence-specific DNA binding ligands. One of these, cysteine cyclic PI polyamides containing two β-alanine molecules, were designed to recognize a 7-bp DNA sequence with high binding affinity. In this study, an efficient cyclization reaction between a cysteine and a chloroacetyl residue was used for dimerization in the synthesis of a unit recognizing symmetrical DNA sequences. To evaluate specific DNA binding properties, dimerized PI polyamide binding was measured using a surface plasmon resonance

(SPR) method. Extending this molecular design, we synthesized a large dimerized PI polyamide that can recognize a 14-bp region in duplex DNA.

Keywords: pyrrole-imidazole polyamide · sequence-specific DNA recognition · selective dimerization · SPR analysis · native PAGE analysis

Introduction

Pyrrole-imidazole (PI) polyamides, which contain *N*-methylpyrrole (Py) and *N*-methylimidazole (Im), bind to the minor groove of DNA with high affinity and specificity.[1] Their DNA recognition depends on a code of side-by-side pairing of Py and Im—a pairing of Im opposite Py targets a G-C base pair, a Py-Im pair targets a C-G base pair, and a Py-Py pair targets an A-T or T-A base pair.[2] To date, various structural motifs of PI polyamides have been studied, such as linear,[3] hairpin,[4] cyclic[5], and tandem[6] motifs. To provide specific functions to PI polyamides, various compounds, such as alkylating agents,[7] fluorescent dyes,[8] and histone deacetylase inhibitors, such as SAHA,[9] have been conjugated to PI polyamides. These functionalized PI polyamide conjugates are expected to show various biological effects on the basis of the binding to their target DNA sequences; their recognition length is 5–9 base pairs for predetermined DNA sequences. Although some linear PI polyamides that recognize long DNA sequences of up to 16 base pairs have been reported,[3l, 4b] the sequence specificity of linear PI polyamides has not been well characterized compared with the conventional

short hairpin PI polyamides. Moreover, the synthesis of longer PI polyamides often suffers from their lower coupling yield.

Recently, we reported the synthesis of cysteine cyclic PI polyamides that were cyclized by an intramolecular reaction between a cysteine and chloroacetyl residue.[10] This intramolecular cyclization reaction was complete within 1 h under mildly basic conditions, and had a high conversion yield without producing any by-product. Some interaction of stacking of aromatic rings and hydrogen bonding of the amide linkage would be expected to induce efficient cyclization of the PI polyamides. To achieve a simple synthesis of PI polyamide with extended base pair DNA recognition, we examined intermolecular dimerization as a method for the synthesis of dimerized cyclic PI polyamides.

Here, we describe a specific dimerized reaction for PI polyamides between cysteine and a chloroacetyl residue. This protocol is simpler than the previous PI polyamide cyclization, and can be applied to longer PI polyamides. We examined the DNA binding affinities and sequence specificities of dimerized cyclic PI polyamides. It was found that some of the cyclic PI polyamides had higher affinities for the target DNA sequence than the mismatch DNA sequences. Using this method, we successfully synthesized a long PI polyamide and showed that it recognized 14 base pairs in a DNA sequence.

Results and Discussion

Synthesis of Dimerized Cyclic PI Polyamide: Previously, a cyclization reaction of PI polyamides using an S_N2 reaction between an *N*-terminal chloroacetyl group and a sulfhydryl group on a cysteine residue was examined. If the coupling reaction proceeds between two monomers, cyclic dimer PI polyamides can be prepared. To test the possibilities, a 4-ring PI polyamide **1** was initially designed for the expected intermolecular reaction between the *N*-terminal chloroacetyl group and the sulfhydryl group on the cysteine. After completion of a machine-assisted Fmoc solid-phase synthesis, the 4-ring PI polyamide **1** was obtained by the cleavage from the resin using TFA (Figure 1). The resulting polyamide **1** was dissolved in basic water conditions (DMF, H_2O, and DIEA) for 30 min, and the progress of the reaction was checked by HPLC analysis. It is notable that the cyclization reaction was completed within only 10 min with high efficiency, as shown in Figure 2 and S1. This reaction time for PI polyamide cyclization was far faster than that of conventional methods. Interestingly, a by-product was observed as a minor peak upon reversed-phase HPLC (Figure S2). These findings showed that the chloroacetyl-cysteine reaction in **1** can be specifically used in dimerized PI polyamide **2** synthesis intermolecularly as well as intramolecularly.[10] Although multimerization of PI polyamide in a cyclization step was reported by Dervan's group,[5e] the reaction reported here provides a significantly higher yield for the dimer formation. After removing solvent *in vacuo*, dimerized cyclic PI polyamide **3** was produced by the coupling with *N*,*N*-dimethyl-1,3-propanediamine

(DMF, DIEA, PyBOP) at room temperature. Following a similar procedure, other types of dimerized cyclic PI polyamides **4–6** were synthesized. Products **3–6** were purified by reversed-phase HPLC, and their chemical structures were confirmed by ^{1}H NMR and electrospray ionization time-of-flight mass spectrometry (ESI-TOF-MS).

DNA Binding Affinities of Dimerized Cyclic PI Polyamides: The binding affinities of **3**, **4**, **5**, and **6** were examined using a surface plasmon resonance (SPR) method. Each PI polyamide was passed to a 5′-biotinylated hairpin DNA immobilized on a sensor chip through a biotin–avidin system. In these experiments, hairpin DNA **III**, **IV**, **V** and **VI** were used—each had a target sequence of **3**, **4**, **5**, and **6**, respectively. In the first trial, the PI polyamides dissolved in HBS-EP buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005 % surfactant P20) with 0.1 % DMSO were used for analysis. Under those conditions, however, an unexpected increase of resonance unit (RU) was observed at high PI polyamide concentration (Figure S3a). The unexpected increase was considered to be caused by the aggregation of the PI polyamides in water. To avoid this aggregation, 2-hydroxypropyl-β-cyclodextrin (HpβCD) was added, which has been reported to increase the water solubility of PI polyamide.[11] With 50 mM HpβCD, the unexpected increase was not observed and so reasonable sensorgrams of each dimerized cyclic PI polyamide were obtained (Figure S3b) and dissociation constants (K_D) could be calculated. Thus, SPR assays were performed in the presence of HpβCD to obtain sensorgrams. The sensorgrams, about **3–6** against each match sequence are shown in

Figure 3 (the other sensorgrams against mismatch sequences are shown in Figure S4-7), and K_D values and the specificities calculated from them are shown in Table 1.

From the K_D values in the matching DNA sequence, their binding affinity was less than that of conventional cyclic PI polyamide (K_D: 10^{-10}–10^{-11}).[5a,5b] In terms of sequence specificity, however, those of these newly synthesized PI polyamides were comparable to previously reported PI polyamides. This is notable because a previous study indicated that some cyclic PI polyamides showed poor sequence specificity[5a] and because these new types of PI polyamides contain a sulfur atom in the bend region. The results suggested that these dimerized cyclic polyamides merit further study as sequence-specific binders.

Large Dimerized Cyclic PI Polyamide: Using the same method as described above, a large, dimerized cyclic PI polyamide **7** containing 10 Py and 10 Im moieties was synthesized successfully. The cyclization step was also completed within 30 min with high conversion efficiency (Figure 4). The product was confirmed by ESI-TOF-MS. This efficiency showed that such a large PI polyamide could be obtained readily by conjugating two PI polyamide molecules.

To investigate the effect of the number of bases corresponding to L-cysteine–β-alanine sequences on binding affinity, a native PAGE assay was performed with DNA **1**, DNA **2**, and DNA **3**, which contain one, two, or three bases, corresponding to L-cysteine–β-alanine sequences, respectively (Figure 5a). In the case of DNA **1** (lanes 1–2, Figure

5b) and **3** (lanes 5–6), no band shift was observed, whereas a clear band shift was observed in the case of DNA **2** (lanes 3–4). These results suggested that two bases was the best number for the L-cysteine–β-alanine sequences of the large cyclic PI polyamide **7**.

The binding and specificity of the large cyclic PI polyamide **7** were also examined using a native PAGE assay. The large cyclic PI polyamide **7** was mixed with match DNA oligomers (DNA **2**, 5′-ATCAGCCTTGGCCTTGGCAGTA-3′, lanes 1–2, Figure 6) or 2-bp mismatch DNA oligomers (DNA **4**, 5′-ATCAG**A**CTTGGCCTTG**A**CAGTA-3′, lanes 3–4). A retarded band was detected in match combination, whereas it was not in mismatch combinations. These results indicated that the large PI polyamide **7** had specificity for its target DNA sequences—a 14-bp recognition sequence, which is the second longest following the 16-bp recognition with a homodimer PI polyamide.[3m] This is the first report of the synthesis and evaluation of PI polyamides containing three components of sequence-recognition sites in a single molecule. Though tandem types of PI polyamide have also been studied to determine extensive sequence recognition, synthesis and specificity problems have been identified.[6a] Thus, the current dimerizing strategy may provide an alternative for synthesizing PI polyamides. By further investigation of this type of PI polyamide synthesis, development of PI polyamides with special selectivity to a certain DNA sequence would be expected; single binding to a human genome without unwanted interactions would be extremely important for clinical use.

Conclusion

In this study, the reaction between a cysteine and a chloroacetyl group was used for an intermolecular reaction in a PI polyamide to specifically synthesize dimerized cyclic PI polyamides. Various types of PI polyamides were synthesized to estimate binding properties to match DNA sequences. Dimerized cyclic PI polyamides were synthesized simply using a new synthetic design under mildly basic conditions, and selective binding to its target DNA sequences, and have good specificities for DNA base pair recognition.

Experimental Section

General: Fmoc-β-wang resin (0.55 mmol/g), HCTU, and Fmoc-Cys (Trt) were purchased from Peptide International. Fmoc-I-COOH, Fmoc-P-COOH, DMF, 1-methyl-2-pyrrolidone (NMP), and piperidine were purchased from Wako, and Fmoc-β-Ala-OH, Fmoc-γ-Abu-COOH, PyBOP were from Novabiochem. Diisopropylethylamine (DIEA), acetic anhydride, and *N*,*N*-dimethyl-1,3-propanediamine (Dp) were from Nacalai Tesque, Inc. Trifluoroacetic acid was from Kanto Chemical Co., Inc. Chloroacetic anhydride was from Aldrich. Biotin modified DNA oligomers were purchased from JBioS. All other reagents and solvents were purchased from standard suppliers and without further purification. SPR assays were performed on a Biacore X system (GE Healthcare), and processing of data was carried out using BIA evaluation, version 4.1. High-performance liquid chromatography (HPLC) analysis was performed with a JASCO PU-2089 Plus HPLC pump, a JASCO UV2075 HPLC UV/VIS detector and a Chemcobond 5-ODS-H reversed phase column (4.6 x 150 mm) in 0.1 % TFA in water with CH_3CN as eluent at a flow rate of 1.0 mL/min, and a linear gradient elution with detection at 254 nm. HPLC purification was performed with a JASCO PU-2089 Plus HPLC pump, a JASCO UV2075 HPLC UV/VIS detector and a Chemcobond 5-ODS-H reversed phase column (10 x 150 mm) in 0.1 % TFA in water with CH_3CN as eluent at a flow rate of 3.0 mL/min, and a linear gradient elution of 20-65 % CH_3CN over 30 min with detection at 254 nm. Collected fractions were analyzed by ESI-TOF-MS (Bruker). Reversed phase

flash chromatography was performed on CombiFlash Rf (Teledyne Isco, Inc.) using a 4.3 gram reversed phase Flash Column (C18 RediSep Rf) in 0.1 % trifluoroacetic acid in water with CH_3CN as the eluent at a flow rate of 18.0 mL/min, and a linear gradient elution of 0 to 35 % CH_3CN in 5 to 30 min with detection at 254 nm. UV spectra were measured on a Nanodrop ND-1000 spectrophotometer.

Fmoc Solid-Phase Syntheses of PI polyamides: Solid phase syntheses were performed on a PSSM-8 (Shimadzu) with a computer-assisted operation system at a 0.05 mmol scale (90 mg of Fmoc-β-wang resin, 0.55 mmol/g) by using Fmoc chemistry. Fmoc unit (0.20 mmol) in each step was set up to solve by NMP on the synthetic line. The following conditions were used in all PI polyamide solid-phase syntheses for each cycle: twice deblocking for 4 min with 20 % piperidine/DMF (0.6 mL), activating for 2 min with HCTU (88 mg, 0.21 mmol) in DMF (1 mL) and 10 % DIEA/DMF (0.4 mL), coupling for 60 min, and washing with DMF. All couplings were carried out with single-couple cycle. Fmoc-PI-CO_2H as a dimer-coupling unit was employed for the difficulty of the coupling a NH_2-I moiety to Fmoc-P-CO_2H. Fmoc-L-Cys (Trt) were employed under the same coupling condition. At the last capping process, the samples were washed for 15 min with 20 % chloroacetic anhydride/DMF. All lines were purged with solution transfers and bubbled by N_2 gas for stirring resin. After the completion of the synthesis, the resin was washed with DMF (2 mL) and methanol (2 mL), and then dried in a desiccator at room temperature in vacuo.

Dp-β-cyclo-((*R*)-Cys-PPIIAc-(*R*)-CysPPIIAc)-β-Dp (3): The resin (120 mg, 0.050 mmol) was placed in a 2-mL plastic vial, trifluoroacetic acid (0.95 mL), triisopropylsilane (0.025 mL), water (0.025 mL) were added, and the solution stirred at room temperature for 1 h. Resin was removed by filtration and washed thoroughly with dichloromethane. The resultant filtrate was removed *in vacuo*. The residue was triturated by diethyl ether, and dried *in vacuo*. The resulting crude (**1**, 12.0 mg) as light-yellow powder was checked by HPLC, ^{1}H NMR and ESI-TOF-Mass (Analytical HPLC: t_R = 11.5 min. ^{1}H NMR [600 MHz, DMSO-d_6] δ 10.71 (s, 1H; NH), 10.32 (s, 1H; NH), 9.95 (s, 1H; NH), 9.44 (s, 1H; NH), 8.03-8.02 (t, *J* = 6.0 Hz, 1H; NH), 7.99-7.98 (d, *J* = 8.0 Hz, 1H; NH), 7.57 (s, 2H; CH), 7.28 (s, 1H; CH), 7.24 (s, 1H; CH), 7.17 (s, 1H; CH), 7.03 (s, 1H; CH), 4.44-4.40 (m, 1H; CH), 4.28 (s, 2H; CH_2), 4.02-3.99 (m, 6H; CH_3), 3.86-3.81 (m, 6H; CH_3), 2.89-2.80 (m, 2H; CH_2), 2.41-2.39 (t, *J* = 6.0 Hz, 2H; CH_2), 2.27-2.24 (t, *J* = 8.0 Hz, 2H; CH_2). ESI-TOF-Mass: *m/z* calcd. for $C_{30}H_{35}ClN_{12}O_8S$: $[M+H]^+$ 759.2188; found: 759.2174). 8.0 mg of the crude was dissolved with DMF (0.6 mL), MilliQ (0.27 mL) and DIEA (0.03 mL). After stirring for 30 min, the peak was consumed and efficiently converted to new peak by HPLC analysis. The cyclic product (**2**) was checked by HPLC, ^{1}H NMR and ESI-TOF-Mass (Analytical HPLC: t_R = 10.4 min. ^{1}H NMR [600 MHz, DMSO-d_6] δ 10.34 (s, 2H; NH), 10.16 (s, 2H; NH), 9.90 (s, 2H; NH), 9.30 (s, 2H; NH), 8.03-8.02 (m, 4H; NH), 7.54 (s, 2H; CH), 7.50 (s, 2H; CH), 7.41 (s, 2H; CH), 7.32 (s, 2H; CH), 6.87 (s, 2H; CH), 6.84 (s, 2H; CH), 4.55 (m, 2H; CH), 4.00 (s, 12H; CH_3), 3.79 (s, 12H; CH_3), 2.61 (m, 8H;

CH_2) 2.38 (m, 8H; CH_2). ESI-TOF-Mass: *m/z* calcd. $C_{60}H_{68}N_{24}O_{16}S_2$ $[M+2H]^{2+}$ 723.2422; found: 723.2438). The solvent was removed *in vacuo*. DMF (0.375 mL), DIEA (0.005 mL), PyBOP (5 mg, 0.01 mmol) and Dp (0.005 ml, 0.05 mmol) were added to a reaction solution at room temperature. The coupling reaction was continuously conducted after stirring for another 1 h, the solvent was removed *in vacuo*. The residue was triturated by diethyl ether and diethylether-dichloromethane, and dried *in vacuo*. The resulting polyamide crude (14.4 mg) was purified by reversed phase flash chromatography. After purification, appropriate fractions were collected under freeze-dry condition to give **3** (2.8 mg, 9 % yield for 9 steps) for the evaluation in DNA binding affinity. (Analytical HPLC: t_R = 9.2 min. ^{1}H NMR (600 MHz, DMSO-d_6) δ10.37 (s, 2H; NH), 10.19 (s, 2H; NH), 9.92 (s, 2H; NH), 9.31 (s, 2H; NH), 8.09-8.02 (m, 6H; NH), 7.56 (s, 2H; CH), 7.51 (s, 2H; CH), 7.40 (s, 2H; CH), 7.30 (s, 2H; CH), 6.90 (s, 4H; CH), 4.52-4.50 (m, 2H; CH), 4.00-3.99 (m, 12H; CH_3), 3.83-3.79 (m, 12H; CH_3), 3.16-3.08 (m, 4H; CH_2), 2.76-2.75 (m, 12H; CH_3), 2.62-2.56 (m, 8H; CH_2), 2.47-2.38 (m, 8H; CH_2), 2.29-2.27 (m, 4H; CH_2), 1.74-1.71 (m, 4H; CH_2). ESI-TOF-Mass: *m/z* calcd. $C_{70}H_{92}N_{28}O_{14}S_2$ $[M+2H]^{2+}$ 807.3473; found 807.3469).

4, **5** and **6** were also synthesized according to the above procedure and purified by reversed phase flash chromatography under similar conditions. The data about these compounds are shown in Supporting Information.

Dp-β-PPI-γ-PII-β-cyclo-((*R*)-Cys-PPIIAc-(*R*)-CysPPIIAc)-IIP-γ-IPP-β-Dp (5): The resin (63.4 mg, 0.022 mmol) was placed in a 2-mL plastic vial, trifluoroacetic acid (0.48 mL), triisopropylsilane (0.0125 mL), water (0.0125 mL) were added, and the solution stirred at room temperature for 1 h. Resin was removed by filtration and washed thoroughly with dichloromethane. The resultant filtrate was removed *in vacuo*. The residue was triturated by diethyl ether, and dried *in vacuo*. The resulting crude (9.2 mg) as light-yellow powder was checked by HPLC and ESI-TOF-Mass (Analytical HPLC: t_R = 11.3 min. ESI-TOF-Mass: *m/z* calcd. for $C_{70}H_{80}ClN_{29}O_{16}S$: $[M+2H]^{2+}$ 825.7952; found: 825.7930). The resulting polyamide crude was purified by reversed phase flash chromatography. After purification, appropriate fractions were collected under freeze-dry condition to give the PI polyamide containing cysteine and a chloroacetyl residue (2.1 mg). The resulting polyamide was dissolved with DMF (0.6 mL), MilliQ (0.27 mL) and DIEA (0.03 mL). After stirring for 30 min, the peak was consumed and efficiently converted to new peak by HPLC analysis. The cyclic product was checked by HPLC and ESI-TOF-Mass (Analytical HPLC: t_R = 9.9 min. ESI-TOF-Mass: *m/z* calcd. $C_{140}H_{158}N_{58}O_{32}S_2$ $[M+2H]^{2+}$ 1614.6059; found 1614.6067). The solvent was removed *in vacuo*. DMF (0.2 mL), DIEA (0.002 mL), PyBOP (2 mg, 0.004 mmol) and Dp (0.003 ml, 0.03 mmol) were added to a reaction solution at room temperature. The coupling reaction was continuously conducted after stirring for another 1.5 h, the solvent was removed *in vacuo*. The residue was triturated by diethyl ether and diethylether-dichloromethane, and dried *in vacuo*. The resulting polyamide crude (2.4

mg) was purified by HPLC. After purification, appropriate fractions were collected under freeze-dry condition to give the large cyclic PI polyamide **7** (0.8 mg, 1 % yield for 17 steps) for the evaluation in DNA binding affinity. (Analytical HPLC: t_R = 9.2 min. ESI-TOF-Mass: *m*/*z* calcd. $C_{150}H_{182}N_{62}O_{30}S_2$ $[M+4H]^{4+}$ 849.8616; found 849.8591.)

SPR Assays: The SPR assays were performed using a BIACORE X instrument. Biotinylated hairpin DNAs were purchased from JBioS (Tokyo, Japan) and the sequences are shown in Tables. Hairpin biotinylated DNA are immobilized to streptavidin-coated sensor chip SA to obtain the desired immobilization level (approximately 900 RU rise). SPR assays were carried out using HBS-EP buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3mM EDTA, and 0.005 % Surfactant P20) with 0.1 % DMSO and 2-hydroxypropyl-β-cyclodextrin (50 mM, for avoiding aggregation) at 25 °C. A series of sample solutions with various concentrations were prepared in the buffer with 0.1 % DMSO and injected at a flow rate of 20 ml/min. To measure the rates of association (k_a), dissociation (k_d) and dissociation constant (K_D), data processing was performed with an appropriate fitting model using the BIAevaluation 4.1 program. The binding model with mass transfer was used for fitting the sensorgrams to give better fitting.

Native PAGE Analysis: Native PAGE assays to evaluate the large cyclic PI polyamide **7** binding were carried out using samples containing 4 μM DNA oligomer (match or 2

bp mismatch) and the large cyclic PI polyamide **7** (0 or 8 μM) with 20 mM cacodylate buffer. Native 20 % polyacrylamide PAGE gels were run at 100 V for 6 h in 1x Tris-borate-EDTA (TBE) buffer (pH 8.3) at 4 °C prior to sample loading. The gels were imaged with Dolphin-View2 (Kurabou).

ACKNOWLEDGMENT

This work was supported by Core Research for Evolutional Science and Technology (CREST) from Japan Science and Technology Agency, a Grant-in Aid for Priority Research from the Minisity of Education, Culture, Sports, Science, and Technology, Japan.

REFERENCE

[1] Trauger, J. W.; Barid, E. E.; Dervan, P. B. *Nature*, **1996**, *382*, 559-561.

[2] a) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Science*, **1994**, *266*, 646-650; b) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.*, **1995**, *117*, 3325-3332; c) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature,* **1998,** *391*, 468-471

[3] a) Pelton, J. G.; Wemmer D. E. *Proc. Natl. Acad. Sci. USA,* **1989**, *86*, 5723-5727; b) Pelton, J. G.; Wemmer D. E. *J. Am. Chem. Soc.*, **1990**, *112*, 1393-1399; c) Fagan, P.; Wemmer, D. E. *J. Am. Chem. Soc.*, **1992**, *114*, 1080-1081; d) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad.*

Sci. USA, **1992**, *89*, 7586-7590; e) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.*, **1992**, *114*, 8783-8794; f) Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochemistry*, **1993**, *32*, 11385-11389; g) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.*, **1993**, *115*, 2572-2576; h) Geierstanger, B. H.; Dwyer, T. J.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.*, **1993**, *115*, 4474-4482; i) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.*, **1993**, *115*, 9892-; j) Dwyer, T. J.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.*, **1993**, *115*, 9900-9906; k) Geierstanger, B. H.; Jacobsen, J. P.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Biochemistry*, **1994**, *33*, 3055-3062; l) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.*, **1998**, *120*, 3534-3535

[4] a) Mrksich, M., Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.*, **1994**, *116*, 7983-7988; b) Trauger, J. W.; Baird, E. E.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.*, **1996**, *118*, 6160-6166; c) de Clairac, R. P.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.*, **1997**, *119*, 7909-7916

[5] a) Cho, J.; Parks, M. E.; Dervan P. B. *Proc. Natl. Acad. Sci. USA*, **1995**, *92*, 10389-10392; b) Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.*,**1999**, *121*, 1121-1129; c) Melander, C.; Herman, D. M.; Dervan P. B. *Chem. Eur. J.* **2000**, *6*, 4487-4497; d) Chenoweth, D. M.; Harki, D. A.; Phillips, J. W.; Dose, C.; Dervan, P. B. *J. Am. Chem. Soc.*,**2009**, *131*, 7182-7188; e) Chenoweth, D. M.; Harki, D. A.; Dervan, P. B. *Org. Lett.* **2009**, *11*, 3590-3593; f) Li, B. C.; Montgomery, D. C.; Puckett, J. W.; Dervan, P. B. *J. Org. Chem.* **2013**, *78*, 124-133

[6] a) Herman, D. M.; Baird, E. E.; Dervan, P. B. *Chem. Eur. J.* **1999**, *5*, 975-983; b) Maeshima, K.; Janssen, S.; Laemmli, U. K. *EMBO J.* **2001**, *20*, 3218-3228; c) Kers, I.; Dervan, P. B. *Bioorg. Med. Chem.*, **2002**, *10*, 3339-3349; d) Schaal, T. D.; Mallet, W. G.; McMinn, D. L.; Nguyen, N. V.; Sopko, M. M.; Jhon, S.; Parekh, B. S. *Nucleic Acids Research,* **2003**, *31*, 1282-1291; e) Sasaki, S.; Minoshima, M.; Fujimoto, J.; Shinohara, K.; Bando, T.; Sugiyama, H. *Nucleic Acids Symposium Series*, **2007**, *51*, 265-266

[7] a) Tao, Z. F.; Fujiwara, T.; Saito, I.; Sugiyama, H. *Angew. Chem. Int. Ed.,* **1999**, *38*, 650-653; b) Tao, Z. F.; Fujiwara, T.; Saito, I.; Sugiyama, H. *J. Am. Chem. Soc.,* **1999**, *121*, 4961-4967; c) Tao, Z. F.; Saito, I.; Sugiyama, H. *J. Am. Chem. Soc.,* **2000**, *122*, 1602-1608; d) Bando, T.; Iida, H.; Saito, I.; Sugiyama, H. *J. Am. Chem. Soc.*, **2001**, *123*, 5158-5159; e) Fujimoto, K.; Iida, H.; Kawakami, M.; Bando, T.; Tao, Z. F.; Sugiyama, H. *Nucleic Acids Research,* **2002**, *30*, 3748-3753; f) Bando, T.; Sasaki, S.; Minoshima, M.; Dohno, C.; Shinohara, K.; Narita, A.; Sugiyama, H., *Bioconjugate Chem*, **2006**, *17*, 715-720; g) Takagaki, T.; Bando, T.; Kitano, M.; Hashiya, K.; Kashiwazaki, G.; Sugiyama. H. *Bioorg. Med. Chem.*, **2011**, *19*, 5896-5902; h) Kashiwazaki, G.; Bando, T.; Yoshidome, T.; Masui, S.; Takagaki, T.; Hashiya, K.; Pandian G. N.; Yasuoka, J.; Akiyoshi, K.; Sugiyama, H. *J. Med. Chem.,* **2012**, *55*, 2057-2066; i) Yoshidome, T.; Endo, M.; Kashiwazaki, G.; Hidaka, K.; Bando, T.; Sugiyama, H. *J. Am. Chem. Soc.,* **2012**, *134*, 4654-4660

[8] a) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U S A.* **2003** *100*, 12063-12068; b) Fechter, E. J.; Olenyuk, B.; Dervan, P. B. *J. Am. Chem.*

Soc., **2005**, *127*, 16685-16691

[9] a) Ohtsuki, A.; Kimura, M. T.; Minoshima, M.; Suzuki, T.; Ikeda, M.; Bando, T.; Nagase, H.; Shinohara, K.; Sugiyama H. *Tetrahedron Lett*. **2009**, *50*, 7288-7292; b) Pandian, G. N.; Ohtsuki, A.; Bando, T.; Sato, S.; Hashiya, K.; Sugiyama, H. *Bioorg*. *Med*. *Chem*., **2012**, *20*, 2656-2660

[10] Morinaga, H.; Bando, T.; Takagaki, T.; Yamamoto, M.; Hashiya, K.; Sugiyama, H. *J*. *Am*. *Chem*. *Soc*., **2011**, *133*, 18924-18930

[11] Hargrove, A. E.; Raskatov, J. A.; Meier, J. L.; Montgomery, D. C.; Dervan, P. B. *J*. *Med*. *Chem*., **2012**, *55*, 5425-5432

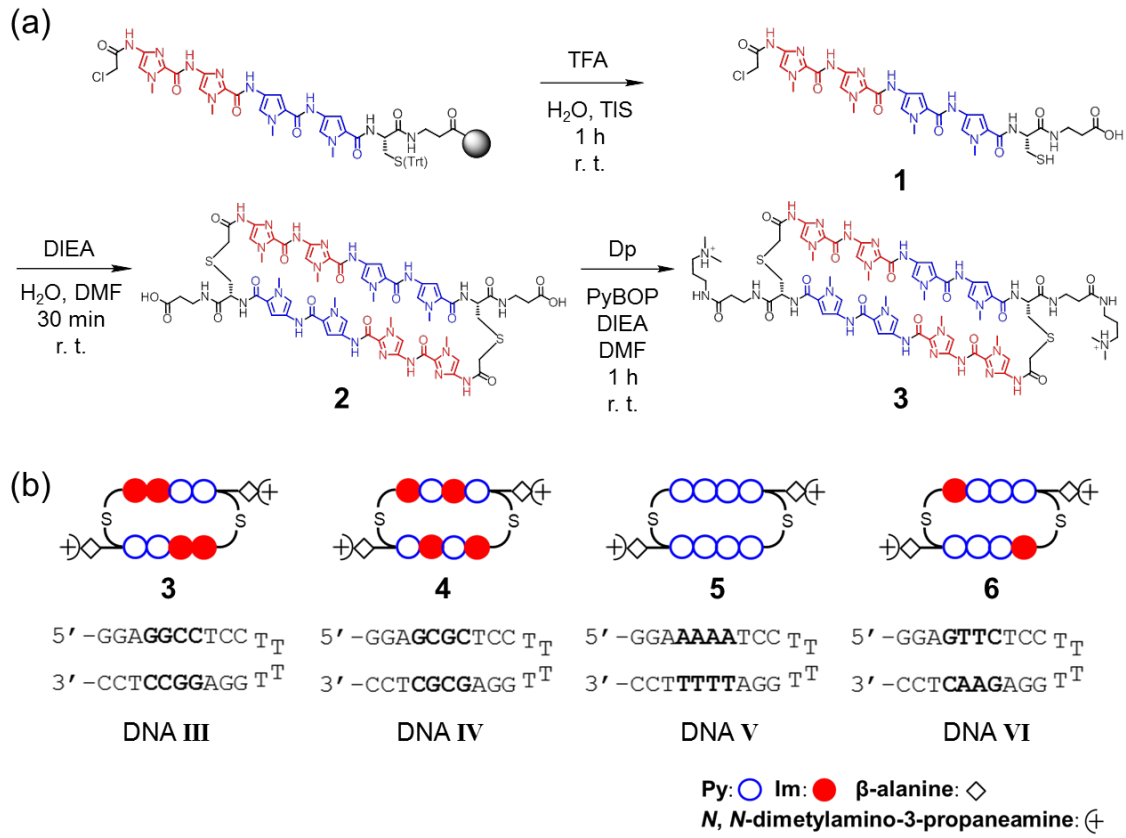


Figure 1. (a) Synthesis of dimerized PI polyamide **3**. (b) The structure of dimerized PI polyamides (**3-6**) and corresponding DNA oligomers containing each match sequence (DNA **III-VI**). The 5′-terminal of each DNA oligomer was actually biotinylated to immobilize them on a streptavidin-coated sensor chip SA for further analysis.

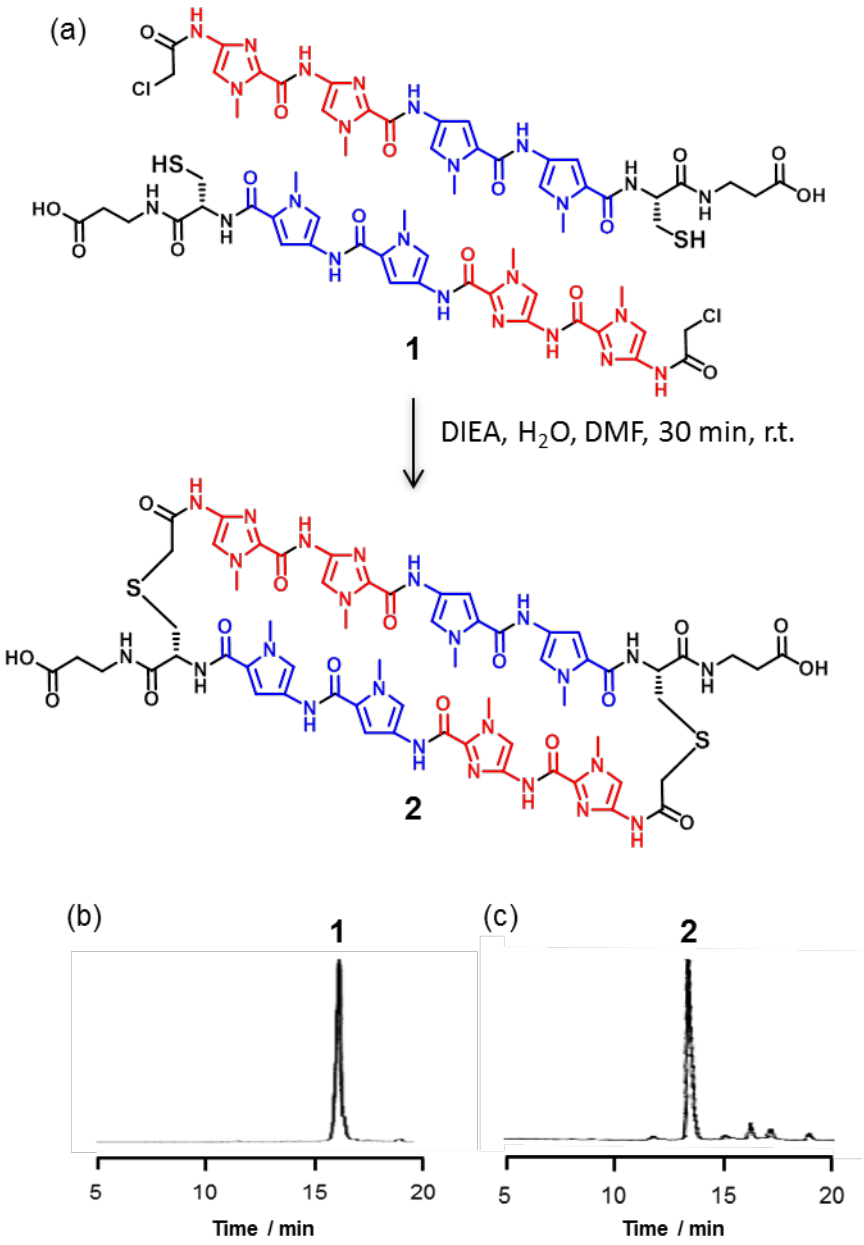


Figure 2. (a) The dimerizing reaction of **1**. The starting material **1** (15.4 min peak) (b) was converted to dimerized product (13.3 min peak). HPLC charts (0.1 % TFA in water with CH_3CN as eluent, and a linear gradient elution of 20 %–80 % CH_3CN over 40 min) before reaction (b), and after reaction (c) are shown.

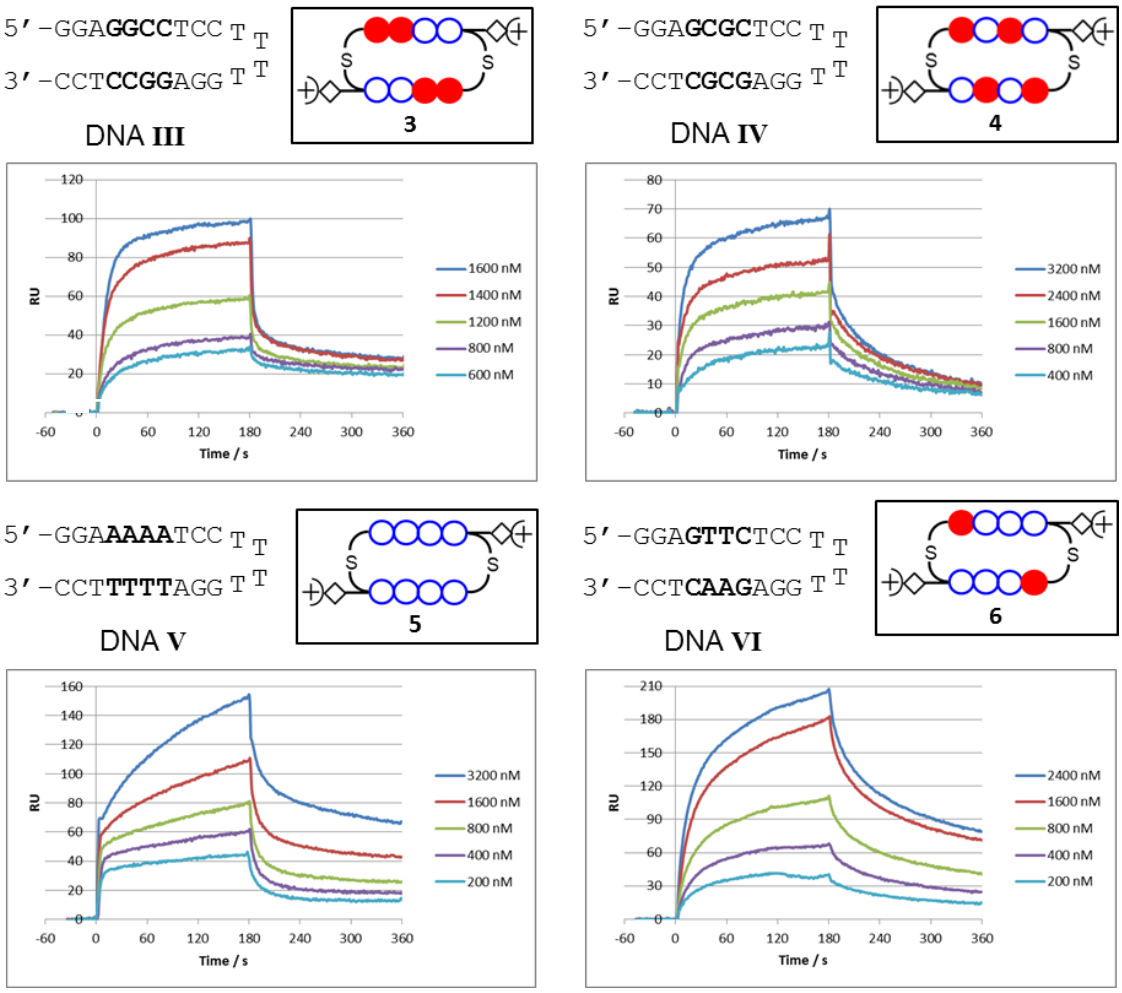
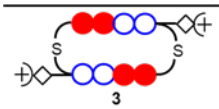
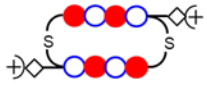
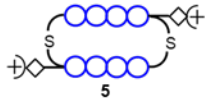
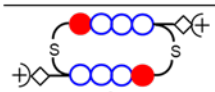


Figure 3. SPR sensorgrams for the dimerized cyclic PI polyamides **3**-**6**. These sensorgrams showed that **3-6** had specificity for each target base pair sequence (DNA **III-VI**).

Table 1. Dissociation constant (K_D) and sequence specificity of four PI polyamides (**3** – **6**) – DNA oligomers combinations [a]

PI polyamide \ DNA sequence	5′-GGA**GGCC**TCC T_T 3′-CCT**CCGG**AGG T^T DNA **III**	5′-GGA**GCGC**TCC T_T 3′-CCT**CGCG**AGG T^T DNA **IV**	5′-GGA**AAAA**TCC T_T 3′-CCT**TTTT**AGG T^T DNA **V**	5′-GGA**GTTC**TCC T_T 3′-CCT**CAAG**AGG T^T DNA **VI**	**specificity**
 **3**	**match** **1.2 x 10^{-7}**	mismatch 3.3 x 10^{-5}	mismatch 1.6 x 10^{-5}	mismatch 1.0 x 10^{-5}	83 ~ 280
 **4**	mismatch 3.7 x 10^{-5}	**match** **5.6 x 10^{-7}**	mismatch 5.0 x 10^{-5}	mismatch 6.4 x 10^{-5}	66 ~ 110
 **5**	mismatch N.D.	mismatch N.D.	**match** **1.1 x 10^{-6}**	mismatch N.D.	> 9.1
 **6**	mismatch 2.7 x 10^{-5}	mismatch 1.7 x 10^{-6}	mismatch 2.6 x 10^{-6}	**match** **7.5 x 10^{-7}**	2.3 ~ 36

[a] Each specificity between match and mismatch DNA sequence is also shown here. The specificity of **5** was determined by supposing N.D. (not determined) as 1.0 x 10^{-5}.

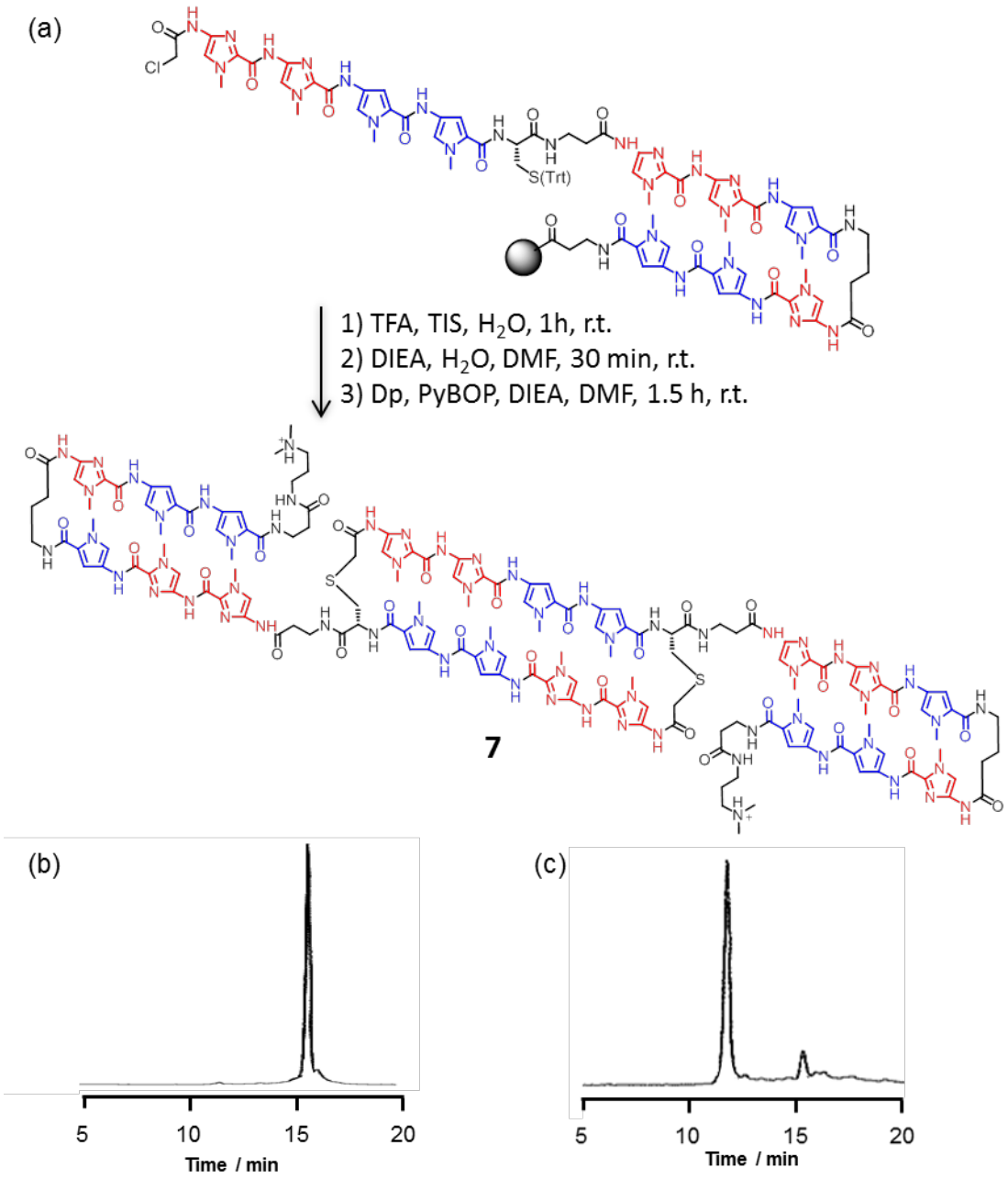


Figure 4. (a) Synthesis of the large dimerized cyclic PI polyamide **7**. Such a large PI polyamide could be synthesized in a simple and quick manner. In particular, cyclization of the PI polyamide was completed in 30 min. The HPLC charts (0.1 % TFA in water with CH_3CN as eluent, and a linear gradient elution of 20 %–65 % CH_3CN over 40 min) before (b) and after (c) cyclization are shown.

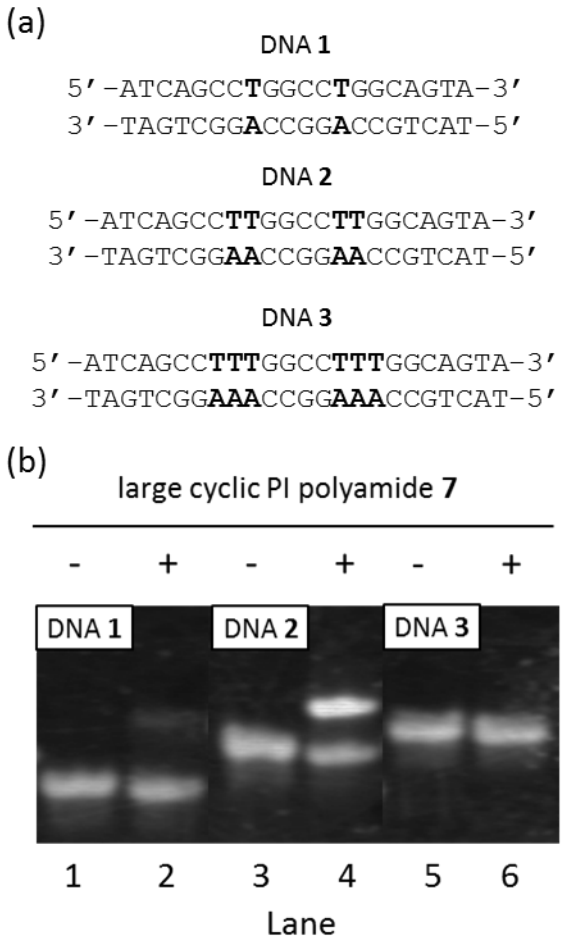


Figure 5. Native PAGE analysis. The proper length of sequences corresponding to the β-alanine–cysteine region. Three types of DNA oligomers that had several number of T/A base pairs (1-3; DNA **1-3**, (a)) as a linker were examined. This finding showed that two T/A linkers were the best match for that β-alanine–cysteine region (b).

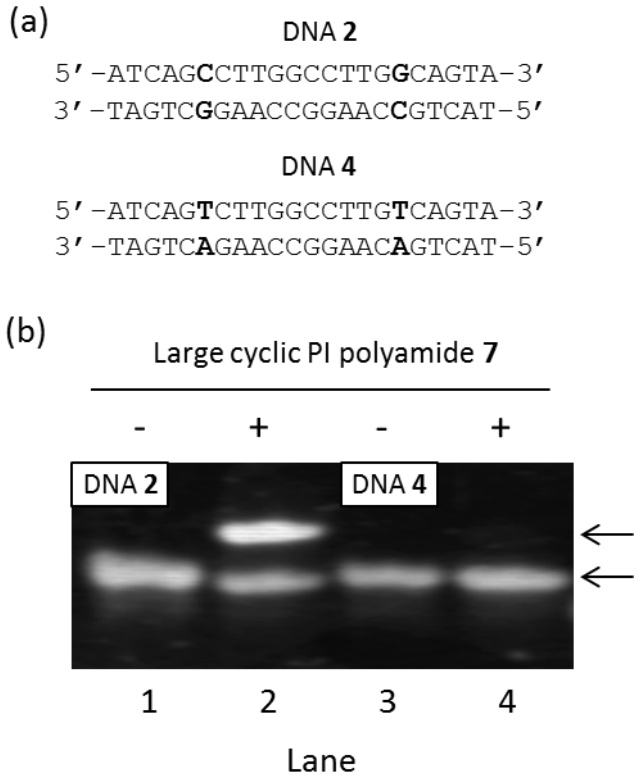


Figure 6. Native PAGE analysis. The specificity of the large cyclic PI polyamide 7 to its target DNA sequences illustrated by native PAGE (b) using match and 2-bp mismatch DNA oligomer (a). The large cyclic PI polyamide **7** was loaded with match DNA oligomers (lane 2) or 2-bp mismatch DNA oligomers (lane 4).

TOC graphic